

Du kommst hier nicht rein: Türsteherfragen identifizieren nachlässige Teilnehmer in Online-Umfragen

Merkle, Belinda; Kaczmirek, Lars; Hellwig, Otto

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Merkle, B., Kaczmirek, L., & Hellwig, O. (2016). Du kommst hier nicht rein: Türsteherfragen identifizieren nachlässige Teilnehmer in Online-Umfragen. *Planung & Analyse : Zeitschrift für Marktforschung und Marketing*, 1, 2-5. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-46660-2>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

planung & analyse

Zeitschrift
für Marktforschung
und Marketing

Eine Marke der
dfv Mediengruppe

SONDERDRUCK

Belinda Merkle, Dr. Lars Kaczmirek, Dr. Otto Hellwig

Du kommst hier nicht rein

„Du kommst hier nicht rein“

TÜRSTEHERFRAGEN identifizieren nachlässige Teilnehmer in Online-Umfragen

B

ei Online-Umfragen gibt es immer die Gefahr von nachlässigen Teilnehmern. Sie kreuzen wahllos an oder wählen immer dieselbe Antwort auf jede Frage. Solche Teilnehmer können die Ergebnisqualität von Online-Umfragen gefährden. Durch sogenannte Türsteherfragen können sie jedoch identifiziert und aussortiert werden. Belinda Merkle, Dr. Lars Kaczmirek und Dr. Otto Hellwig zeigen Best Practice Beispiele, wie solche Qualitäts-Check-Fragen passend zur Umfrage gestaltet werden können.

Wahrheitsgemäße Antworten sind eine wesentliche Bedingung für das Gelingen einer Umfrage. Teilnehmer, die diese Bedingung nicht erfüllen, erzeugen nicht nur zusätzliche Kosten, sondern können vor allem auch die Validität einer Umfrage gefährden. Nun gibt es immer wieder nachlässige Teilnehmer, die einzelne oder mehrere Fragen vorsätzlich oder auch ohne böse Absicht so unaufmerksam bearbeiten,

dass deren Antworten nicht mit ihrer wahren Meinung übereinstimmt oder nicht den Lebensumständen entspricht. Als Ergebnis erhält der Forscher verschiedene Antworten, die kaum mit der Fragestellung zusammenhängen. Zur Gruppe nachlässiger Teilnehmer gehört beispielsweise der so genannte *Random Responder*, welcher wahllos irgendwelche Antworten ankreuzt, und der *Straightliner*, der in einer Fragebatterie von oben bis unten immer die gleiche Antwort wählt. Es gibt auch nachlässige Teilnehmer, die nur oberflächlich lesen und dadurch ungenaue Antworten geben. Was jeweils als „zu nachlässige Teilnahme“ gilt, hängt von der Fragestellung und dem Forschungsziel ab.

Somit ist es wichtig, umfragespezifisch zu entscheiden, welche Teilnehmer für die jeweilige Umfrage zu nachlässig sind, besonders wenn die Erreichung der Zielquoten an Teilnehmern einen weiteren wichtigen Faktor für das Gelingen einer Umfrage darstellt. Das bedeutet, dass ein Cut-Off-Wert für Nachlässigkeit für jede Umfrage

neu definiert werden muss. Teilnehmer, die diesen Cut-Off-Wert überschreiten, sollten dann für den Panel-Anbieter und Forscher möglichst frühzeitig identifiziert und vor Beginn einer Umfrage ausgescreent werden. Während gleichzeitig „nicht-perfekte“ Teilnehmer, welche die Forschungsfrage jedoch nicht gefährden, beibehalten werden können. Ein solcher Cut-Off-Wert kann über das Bestehen oder Nicht-Bestehen einer Qualitäts-Check-Frage operationalisiert werden.



Identifizierung von nachlässigen Teilnehmern

Zur Identifizierung von nachlässigen Teilnehmern mithilfe von Qualitäts-Check-Fragen stellen wir vier Arten von Qualitäts-Check-Fragen vor. Jede Frage soll überprüfen, ob Teilnehmer gewissenhaft antworten. Allerdings zielen sie jeweils auf unterschiedliche Bereiche innerhalb einer Frage ab. Um zu prüfen, ob Teilnehmer Instruktionen auch lesen, stellten Oppenheimer,

Der Umfrager 1

Kaufentscheidung

Im Folgenden werden Ihnen zwei Produkte präsentiert. Bitte lesen Sie sich die Informationen zu beiden Produkten sorgfältig durch und beantworten Sie im Anschluss die Fragen für beide Produkte.



Produkt A

Das Smartphone A ist zu einem Preis von 500 Euro erhältlich ohne Vertrag. Es enthält eine Kamera mit 12 Megapixeln und hat ein 6 Zoll großes Display. Verfügbar ist es in den Farben blau, weiß und schwarz.



Produkt B

Das Smartphone B ist zu einem Preis von 400 Euro erhältlich ohne Vertrag. Wir interessieren uns besonders dafür, ob Sie wirklich die Zeit verwenden und die Einführung sorgfältig lesen, denn wenn Sie dies nicht tun, ist unsere Umfrage leider nicht mehr zuverlässig. Darum bitten wir Sie, um zu zeigen, dass Sie den Einführungstext gelesen haben, die folgenden Antworten sowie den „Weiter“ Button zu ignorieren und dafür auf den Titel „Kaufentscheidung“ zu klicken. Vielen Dank für Ihr Verständnis! Verfügbar ist es in den Farben blau, weiß und schwarz.

Welches Smartphone würden Sie kaufen?

- ☐ Smartphone A ☐ Smartphone B

1

Werden Instruktionen vollständig gelesen? **Instructional Manipulation Checks:** Dabei werden die Teilnehmer in den Instruktionen aufgefordert, die Antwortoptionen zu ignorieren und stattdessen mit einer anderen Aktion zu bestätigen, dass sie die Instruktionen vollständig gelesen haben.

2

Wird eine Frage vollständig gelesen und verstanden? **Instructed Response Items:** Solche Qualitäts-Check-Fragen bestehen aus der Aufforderung, eine bestimmte Antwortoption anzukreuzen.

Der Umfrager 2

Eingabemedium

Antworten Sie mit „Ich stimme vollkommen zu“ auf dieses Item.



- ☒ Ich stimme überhaupt nicht zu
- ☐ Ich stimme nicht zu
- ☐ Ich bin neutral
- ☐ Ich stimme zu
- ☐ Ich stimme vollkommen zu

Der Umfrager 3

Einkommensarten

Es folgt eine Liste mit verschiedenen Einkommensarten, die ein Haushalt beziehen kann. Bitte geben Sie alle Einkommensarten an, die auf Ihren Haushalt zutreffen (Mehrfachantworten sind erlaubt)!

- ☒ Lohn und Gehalt (auch von Auszubildenden)
- ☐ Einkommen aus selbstständiger Tätigkeit (Selbstständige, Landwirte, freiberuflich Tätige)
- ☐ Regelmäßig private Unterhaltszahlungen (z.B. für Geschiedene oder Kinder)
- ☒ Um zu zeigen, dass Sie aufmerksam sind, bitte diese Antwort zusätzlich ankreuzen
- ☐ Arbeitslosengeld I, (Saison-) Kurzarbeitergeld u.ä.

3

Werden alle Antworten gesehen und gelesen? **Instructed Response Items for check-all-that-apply questions:** Diese gleichen den Instructed Response Items mit dem einzigen Unterschied, dass die spezielle Antwortinstruktion nicht in der Fragestellung steht, sondern erst innerhalb einer der verschiedenen Antwortoptionen.



Wahrheitsgemäße Antworten sind Bedingung für das Gelingen einer Umfrage



Meyvis und Davidenko *Instructional Manipulation Checks* vor. Hierbei werden Ausfüllanweisungen in den Instruktionen versteckt. Um zu testen, ob Teilnehmer die Frage lesen, stehen sowohl *Bogus Items* als auch *Instructed Response Items* zur Verfügung. Diese wurden von Beach sowie Meade und Craig beschrieben. Erstere sind Scheinfragen mit einer eindeutig richtigen Antwort, während bei letzterem eine Ausfüllanweisung im Fragentext gegeben wird. Für Mehrfachwahlantworten schlagen wir vor, die Ausfüllanweisung auch in den Antwortoptionen zu verstecken (*Instructed Response Items for check-all-that-apply questions*).



Werden Instruktionen vollständig gelesen?

Instruktionen bilden häufig den Anfang eines Fragebogens können aber auch spezifisch einzelnen Items vorangestellt werden. Je nach Untersuchungsziel kann das vollständige Lesen und Umsetzen der Instruktionen beziehungsweise Einführungstexte von entscheidender Bedeutung sein. Besonders bei der Präsentation verschiedener Produkte kann es wichtig sein, dass sich die Teilnehmer die Einführungstexte genau durchlesen. Um Teilnehmer auszufiltern, die diesen Anforderungen nicht gerecht werden, kommen *Instructional Manipulation Checks* zum Einsatz (siehe Abbildung 1). Dabei werden die Teilnehmer in den Instruktionen aufgefordert, die Antwortoptionen zu ignorieren und stattdessen mit einer anderen Aktion zu bestätigen, dass sie die Instruktionen vollständig gelesen haben. Diese Aktion kann etwa das Klicken auf eine Überschrift oder auf ein anderes Objekt auf dem Bildschirm sein.

Oppenheimer et al. belegen, dass das Ausfiltern von Teilnehmern mithilfe von *Instructional Manipulation Checks* die Ergebnisse einer Studie signifikant verbessert. Dies ist möglich, weil nachlässige Teilnehmer durch ihr Antwortverhalten zusätzliches Rauschen in den Daten erzeugen, das unabhängig von den Fragen ist und damit die Teststärke von statistischen Verfahren reduziert. Allerdings sollte beachtet werden, dass nicht alle Instruktionen gleich wichtig sind. So ist das Ignorieren einer Überleitung weniger dramatisch als das Nicht-Beachten einer komplizierten Definition oder einer Produktdarstellung (zum Beispiel bei Vignettendesigns). Daher ist es wichtig, dass die Ausfüllanweisung in einem für die Studie relevanten Teil der Instruktionen gegeben wird, um zu verhindern, dass voll kooperierende, aber etwas eilige Teilnehmer ausgefiltert werden.

Ebenso ist es gute Praxis Instruktionen insgesamt sparsam zu verwenden und sie dort einzusetzen, wo sie direkt benötigt werden.



Wird eine Frage vollständig gelesen und verstanden?

Um zu testen, ob Teilnehmer die Fragen auch tatsächlich lesen und verstehen, gibt es zwei verschiedene Arten von Qualitäts-Check-Fragen: Scheinfragen (*Bogus Items*) sowie Fragen mit vorgegebener Antwort (*Instructed Response Items*).

Bei Scheinfragen ist die korrekte Antwort bereits bekannt. Sie erfordert somit von der Versuchsperson das vollständige Lesen und Verstehen der Frage. Ein Beispiel ist die Frage „Benutzen Sie gerade ein elektronisches Gerät zum Ausfüllen der Umfrage?“. Sie wird als bestanden gewertet, wenn Teilnehmer einer Online-Umfrage entweder mit „ich stimme zu“ oder mit „ich stimme vollkommen zu“ antworten. Meade und Craig haben solche Scheinfragen mit anderen Methoden verglichen und nachgewiesen, dass sie geeignet sind, um nachlässige Teilnehmer zu entdecken. Ein Nachteil dieser Scheinfragen ist allerdings, dass nicht für jede Umfrage dieselbe Frage universell anwendbar ist. Zudem kann es bei schlecht konstruierten Scheinfragen vorkommen, dass zweideutige Interpretationen möglich sind. Außerdem diskutieren Meade und Craig die Gefahr, dass einzelne Teilnehmer die Scheinfrage für so unsinnig oder lustig halten, dass sie absichtlich eine falsche Antwort auswählen.

Die einfachste Methode, um unachtsame Teilnehmer auszuschließen und die genannten Hürden zu umschiffen, sind daher *Instructed Response Items*. Solche Qualitäts-Check-Fragen bestehen aus der Anforderung, eine bestimmte Antwortoption anzukreuzen (Abbildung 2).

Von Nachteil ist allerdings, dass diese Qualitäts-Check-Fragen leichter von Teilnehmern als Testfragen zu entdecken sind. Dafür bieten sie den Vorteil, dass sie einfacher zu erstellen und universell für jede Umfrage anwendbar sind. Zudem sind keine falschen Interpretationen möglich, weder auf Seiten der Teilnehmer noch auf Sei-

ten der auswertenden Person. Weil es nur eine richtige Antwort gibt, haben *Instructed Response Items* eine höhere Wahrscheinlichkeit nachlässige Teilnehmer zu entdecken als Scheinfragen.



Werden alle Antworten gesehen und gelesen?

Schließlich kann bei einer Frage auch das sorgfältige Beachten aller Antwortoptionen im Fokus liegen, besonders, wenn die Möglichkeit zu Mehrfachwahloptionen gegeben ist. In diesem Fall kann unvollständiges Lesen die Umfrage gefährden, wenn die oberen Antworten bereits sinnvoll erscheinen und Teilnehmer dann nicht die vollständige Liste aller möglichen Antworten lesen, bevor sie sich entscheiden. Um auch solche Problemfälle zu identifizieren, empfehlen wir die Verwendung von *Instructed Response Items for check-all-that-apply questions*. Diese gleichen den *Instructed Response Items* mit dem einzigen Unterschied, dass die spezielle Antwortinstruktion nicht in der Fragestellung steht, sondern erst innerhalb einer der verschiedenen Antwortoptionen (siehe Abbildung 3).

Es gibt Fragen, bei denen es nicht notwendig ist, alle Antworten zu lesen. Wenn es nur eine eindeutig richtige Antwort gibt, ist es aus Teilnehmerperspektive sinnvoll, nicht auch noch die anderen Antworten zu lesen, wenn man seine Antwort bereits gefunden hat. Es gibt viele Teilnehmer, die dann ohne weitere Umschweife mit der nächsten Frage weitermachen. Die Anwendung der *Instructed Response Items for check-all-that-apply questions* in den Antworten bei Einfachauswahl-Fragen würde demnach zum Ausfiltern von kooperierenden Teilnehmern führen, weil es sowohl die richtige Antwort gibt und zusätzlich noch die Instruktion in einer anderen Antwort. Es ist daher zu beachten, dass diese Items nur für Fragebögen mit Mehrfachwahlitems angewendet werden.



Teilnehmer einfach mal wachrütteln

Es existiert noch ein anderer Ansatz, der Qualitäts-Check-Fragen nicht zum Ausschließen von nachlässigen Teilnehmern nutzt, sondern vielmehr zum Wachrütteln, um die Aufmerksamkeit der Teilnehmer zu verbessern. In solch einer Umsetzung können die Teilnehmer erst mit der Umfrage fortfahren, wenn sie die Qualitäts-Check-Frage richtig beantwortet haben. Hier stellt sich die Forschungsfrage, ob man mit solch einem womöglich als aggressiv erlebten Vorgehen Teilnehmer demotiviert. Oppen-

heimer, Meyvis und Davidenko fanden bei Instructional Manipulation Checks jedoch eine positive Wirkung. Sie empfehlen den Einsatz zu Beginn einer Umfrage, um die Gewissenhaftigkeit bei den Teilnehmern sicherzustellen.

Kritik an Qualitäts-Check-Fragen

Häufig diskutierte gemeinsame Nachteile aller Qualitäts-Check-Fragen sind die mögliche Reduzierung der Motivation und Kooperation von Teilnehmern, da sich diese „überwacht“, nicht ernstgenommen, gekränkt fühlen könnten oder ganz einfach die Sinnhaftigkeit der Frage anzweifeln.

Beim Einsatz von Qualitäts-Check-Fragen nimmt man an, dass nachlässiges Verhalten ein generelles Verhalten in der jeweiligen Umfrage ist und nicht nur für einzelne Fragen gilt. Man generalisiert also das Ergebnis einzelner Qualitäts-Check-Fragen auf alle vergangenen oder zukünftigen Antworten des Teilnehmers. Wenn ein Teilnehmer beispielsweise den ersten Check besteht, aber einen zweiten nicht, so würde man den gesamten Datensatz von der Analyse ausschließen. Diese Methoden eignen sich daher vor allem für den Einsatz zu Beginn von Umfragen für Screenouts, jedoch schon aus ethischen Überlegungen weniger für Screenouts im weiteren Verlauf von Umfragen. Bei Checks inmitten oder sogar gegen Ende einer Umfrage sollte man Teilnehmer nicht ausscreenen, sondern lediglich im Nachgang zur Umfrage von der Analyse ausschließen. Situative Nachlässigkeit, die bei einzelnen (schwierigen oder schlechten) Fragen entsteht, kann damit selbstverständlich nicht identifiziert werden.

Die genannten Autoren haben die hier dargestellten Methoden zum Einsatz von Qualitäts-Check-Fragen anhand von Stichproben mit Studierenden an zwei amerikanischen Universitäten geprüft. Wie hoch der Anteil an fälschlich ausgeschlossenen Teilnehmern ist, wenn man solche Fragen in Deutschland anwendet, ist nicht geklärt. In den Versuchen hielt sich die Kritik von Teilnehmern an solchen Fragen zwar in Grenzen. Allerdings wird hierzulande die Verwendung solcher „nicht-echten“ Fragen unter Forschern kritisch diskutiert. Manche bevorzugen weniger sichtbare Kriterien zur Qualitätssicherung wie die Antwortzeiten und die Analyse von Antwortmustern. Wenn man annimmt, dass die große Mehrheit an Teilnehmern zufriedenstellend engagiert an Umfragen teilnimmt, so besteht die Gefahr, diese größere Gruppe mit für sie scheinbar überflüssigen (Trick-)Fragen zu verunsichern

oder zu verärgern. Dies kann dann dazu führen, dass die Qualitäts-Check-Fragen nicht korrekt bearbeitet werden und man dadurch vorher motivierte Teilnehmer verliert. Selbstverständlich besteht auch bei dieser Methode – wie fast immer, wenn man während einer Studie Teilnehmer ausschließt – die Gefahr von Ergebnisverzerrungen, wenn sich die nachlässigen Teilnehmer in ihren wahren Einstellungen und Eigenschaften von der teilnehmenden Gruppe systematisch unterscheiden. Oppenheimer et al. untersuchten diese möglichen Ergebnisverzerrungen. Abgesehen von Motivationsunterschieden sowie Unterschieden in der Tendenz nachlässig zu antworten, fanden sie keine signifikanten Unterschiede zwischen den Teilnehmern, welche die Instructional Manipulation Checks beim ersten Versuch bestehen, und den Teilnehmern, die zunächst eine falsche Antwort geben. Auch Hinweise auf eine Kränkung der Teilnehmer waren nicht zu finden.

Empfehlungen für Türsteherfragen

Aufgrund der Literatur empfehlen wir die Qualitäts-Check-Fragen in Format, Aussehen und Inhalt weitestgehend an echte Fragen anzupassen. Im Durchschnitt sollte pro 50 bis 100 Items eine Qualitäts-Check-Frage gestellt werden und nicht mehr als drei in einer Umfrage. Die Qualitäts-Check-Fragen sollten gemäß den spezifischen Anforderungen der Umfrage ausgewählt werden, umso nachlässige Teilnehmer zu identifizieren.

Für einen massenhaften standardisierten Einsatz in jeder Umfrage sind die vorgestellten Verfahren unserer Ansicht nach nicht geeignet. Sie können gemeinsam mit etablierten Qualitätsmaßen verwendet werden.

Da das Lesen und Verstehen von Fragen für jede Umfrage von Bedeutung ist, empfehlen wir die Verwendung eines Instructed Response Items. Es sollte außerdem eine weitere Qualitäts-Check-Frage hinzugefügt werden, um die Wahrscheinlichkeit zu erhöhen, dass nachlässige Teilnehmer entdeckt werden. Diese sollte dann an die spezifischen Anforderungen der Umfrage angepasst werden, also beispielsweise bei Fokus auf Instruktionen ein Instructional Manipulation Check. Liegt kein besonderer Fokus vor, empfehlen wir ein weiteres Instructed Response Item zu verwenden, um nicht versehentlich mit einer der anderen beiden spezifischeren Methoden in dieser Umfrage brauchbare Teilnehmer auszufiltern.

Die Autoren



Belinda Merkle studiert Psychologie an der Universität Mannheim und an der University of Florida. Sie war Praktikantin bei GESIS – Leibniz Institut für Sozialwissenschaften und in der Arbeitsgruppe für Methodenlehre und Diagnostik der Universität Mannheim.

belinda.merkle@gmx.de



Dr. Lars Kaczmarek ist Leiter des Teams Survey Programs bei GESIS – Leibniz-Institut für Sozialwissenschaften. Er hat das Programmkomitee der General Online Research Konferenz geleitet und war Vorstandsmitglied der deutschen Gesellschaft für Online Forschung.

lars.kaczmarek@gesis.org



Dr. Otto Hellwig hat am Institut für angewandte Sozialforschung der Universität Köln geforscht und promoviert. 2005 gründete er die respondi AG. Seit 2013 ist er Vorstandsvorsitzender der Deutschen Gesellschaft für Online-Forschung (DGOF).

otto.hellwig@respondi.com



Literatur

- Beach, D. A. (1989): *Identifying the random responder. The Journal of psychology*, 123(1), S. 101-103.
- Meade, A. W.; Craig, S. B. (2012): *Identifying careless responses in survey data. Psychological methods*, 17(3), S. 437.
- Oppenheimer, D. M.; Meyvis, T.; Davidenko, N. (2009): *Instructional manipulation checks: Detecting satisficing to increase statistical power. Journal of Experimental Social Psychology*, 45(4), S. 867-872.